

Guaranteed sensor coverage with the *weighted- D^2* sampling

Ajay Deshpande^a,

^aIBM T. J. Watson Research Center, Yorktown Heights, NY, USA

Abstract

In this paper we focus on the mobile sensor coverage problem formulated as a continuous locational optimization problem. Cortès et al. first proposed a distributed version of the Lloyd descent algorithm with guaranteed convergence to a local optima. Since then researchers have studied a number of variations of the coverage problem. The quality of the final solution with the Lloyd descent depends on the initial sensor configuration. Inspired by the recent results on a related k -means problem, in this paper we propose the *weighted- D^2* sampling to choose the initial sensor configuration and show that it yields $O(\log k)$ -competitive sensor coverage before even applying the Lloyd descent. Through extensive numerical simulations, we show that the initial coverage with the *weighted- D^2* sampling is significantly lower than that with the uniform random initial sensor configuration. We also show that the average distance traveled by the sensors to reach the final configuration through the Lloyd descent is also significantly lower than that with the uniform random configuration. This also implies considerable savings in the energy spent by the sensors during motion and faster convergence.

1 Introduction

Sensor coverage is a fundamental issue in large-scale sensing. It addresses the question of when and where to place the sensors. With the advent of miniaturized wireless sensors, sensing is now possible at unprecedented scales. Naturally, sensor coverage has received significant attention in the last decade and to date many different formulations have been proposed. In all the formulations, sensor coverage formalizes a metric for the *quality-of-sensing* and the sensor coverage problem aims to find an optimal or a near-optimal sensor placement which optimizes the metric.

In this paper we focus on a sensor coverage problem that is formulated as a *locational optimization problem* as in [2, 4]. Locational optimization problems have long been studied in various fields in the context of static spatial resource allocation and solved in a centralized computing environment [2, 10]. Cortès et al. for the first time formulated the mobile sensor coverage problem as a locational optimization problem and developed a framework for distributed control and coordination of mobile sensors to obtain optimal coverage [2, 4]. A *centroidal Voronoi configuration*, where each sensor is at the centroid of its Voronoi partition, is well-known to be an optimal solution to this problem [10]. Cortès et al. proposed a distributed *Lloyd gradient descent* algorithm with the guaranteed convergence to a centroidal Voronoi configuration. According to this control algorithm, starting

with an initial configuration each sensor incrementally moves towards the centroid of its Voronoi partition until the convergence [2, 4]. This work initiated the study of several variations of the mobile sensor coverage problem, for example, coverage with limited range constraints for sensing and communication [3], coverage with a generalized monotonic sensing function [12], coverage with sensor consensus and learning [13], and coverage with sensors having location-dependent sensing performance [6]. Every work focused on developing a distributed control law for sensors movement, typically as an extension of the Lloyd descent algorithm, and proved the guaranteed convergence to a locally optimal solution.

An important consideration that is seldom addressed in this area is the quality of the (locally) optimal solution obtained with the Lloyd descent. It depends on the initial sensor configuration. In the works so far, the initial sensor configuration is typically chosen uniformly at random. There is no guarantee on the optimal solution thus obtained. In this paper, we propose *weighted- D^2* sampling to choose the initial sensor configuration, which guarantees $O(\log k)$ -competitive sensor coverage before even applying the Lloyd descent algorithm. Here, k denotes the number of sensors. The application of the Lloyd descent only improves the sensor coverage. Our work is inspired based on the recent work [1] by Arthur and Vassilvitskii on using D^2 sampling for choosing the initial centers in the related *discrete k -means problem* for clustering with the $O(\log k)$ -competitive guarantee. We prove our result in two steps. In the first step, we establish a close relationship between the sensor cover-

Email address: ajayd@us.ibm.com (Ajay Deshpande).

age problem and a suitably selected discrete weighted k -means problem. In the second step, we extend the original result in [1] to the weighted- D^2 sampling for the weighted k -means problem. To the best of our knowledge, this is a first foray into the performance guarantees for sensor coverage before applying any control algorithm. In [11], the authors proposed a deterministic annealing approach adapted to the locational optimization problem to obtain a globally optimal solution. However, there is no guarantee on the convergence rate.

The weighted- D^2 sampling is computationally straightforward. We first create a discrete set of candidate sensor locations using a grid and assign a weight to each location. Then we choose the first sensor location at random from the discrete candidate set with the probability proportional to the weight. Subsequently, we iteratively select the next sensor location with the probability proportional to the weight times the square of the shortest distance to the sensor locations already chosen. Through extensive numerical simulations, we observe that with the weighted- D^2 sampling, the initial sensor coverage before applying the Lloyd descent is significantly lower than that with the uniform random deployment.

Another advantage of the weighted- D^2 sampling for the initial configuration is the energy spent by mobile sensors moving during the Lloyd descent. While recent works (e.g. [5, 9]) aim to develop energy balancing Lloyd descent control laws, this applies once the initial configuration is chosen. The weighted- D^2 sampling already with the guarantee on the initial sensor coverage helps in significantly lowering the energy spent during the descent. Through our simulations, we find that with the weighted- D^2 sampling, the average distance traveled by the sensors to the convergence during the Lloyd descent is considerably lower than that with the uniform random initial configuration. This also means that the energy spent is also significantly lower and depending on the sensor motion dynamics, the convergence is faster.

Organization: In Section 2 we discuss the related work to the sensor coverage and the k -means problems. In Section 3 we define our sensor coverage problem, discuss the Lloyd descent and propose the weighted- D^2 sampling method for choosing the initial sensor configuration. In Section 4, we establish the relationship between the sensor coverage and the discrete weighted k -means problem. In Section 5 we prove the guarantees on the solution to the weighted k -means problem with the weighted- D^2 sampling. In Section 6, we show present results of the numerical simulations. In Section 7 we conclude and outline future work.

2 Related work

Locational optimization problems have long been studied in the areas of spatial economics and facility loca-

tion [2, 10]. Okabe and Suzuki present a review of a class of continuous locational optimization problems which can be addressed using Voronoi diagrams [10]. Cortès et al. formulated the mobile sensor coverage problem as a continuous locational optimization problem for the first time [2, 4]. We will discuss their formulation in detail in the next section. They proposed a

Since then researchers have studied several variations of the mobile sensor coverage. Examples include coverage using sensors with limited-range sensing and communication [3], coverage with a generalized monotonic sensing function [12], coverage with sensor consensus and learning [13], and coverage over networks using sensors with location-dependent sensing performance [6]. All these works have developed variations of the Lloyd descent algorithm and proved guaranteed convergence to local optimal in a distributed setting. In recent years, there has also been some work on developing control laws that balance coverage as well as energy across sensors [5, 9]. In all these works, the initial sensor placement is chosen uniformly at random. In [7] Deshpande et al. developed a version of the Lloyd descent algorithm for the sensor placement in physical networks such as water distribution networks. Using simulations, they showed that the initial sensor configuration with the D^2 sampling yields an optimal solution with better coverage than the uniform random placement, converges faster, and is also significantly closer to the globally optimal solution. In all these works there is no guarantee on the quality of the optimal solution. In [11], the authors proposed a deterministic annealing approach adapted to a class of locational optimization problems to obtain globally optimal solutions. However, there is no guarantee on the convergence rate.

A related problem in the discrete world is the k -means problem which aims to choose k centers for a given set of points such that the sum of the squared distance between each point and its closest center is minimized. This naturally partitions the points into k clusters with the points in the same cluster being the closest to the same center. The center for each cluster turns out to be the centroid of the points in that cluster. k -means is widely used as a clustering technique in unsupervised learning. While k -means problem is known to be *NP-hard*, a locally optimal solution can be obtained very fast using the Lloyd descent algorithm. In [1], Arthur and Vassilvitskii introduce D^2 -sampling to choose k initial points and show that the solution is already $O(\log k)$ -competitive. We will discuss D^2 -sampling in detail in this paper.

3 Coverage Problem and weighted- D^2 sampling

In this section we first introduce the sensor coverage problem and discuss the Lloyd descent algorithm. Then we propose the weighted- D^2 sampling.

3.1 Coverage problem formulation

We consider the same coverage problem formulation in 2D as in [2]. Let Q denote a bounded convex region in \mathbb{R}^2 . Let $\phi(\cdot)$ denote a nonnegative scalar density function $\phi : Q \rightarrow \mathbb{R}_+$. $\phi(q)$ can be viewed as a measure of importance of covering location q . Without loss of generality we assume that $\int_Q \phi(q) dq = 1$. Let $P = (p_1, p_2, \dots, p_k)$

denote the locations of k mobile sensors in Q . The sensing performance at location q due to the i th sensor located at p_i degrades with the Euclidean distance $\|q - p_i\|$ between q and p_i . We assume that it is modeled as $\|q - p_i\|^2$. For fixed sensor locations (p_1, p_2, \dots, p_k) , the sensing performance function induces a *Voronoi partition* $\mathcal{V}(P) = \{V_1, V_2, \dots, V_k\}$ of Q , where,

$$V_i = \{q \mid \|q - p_i\|^2 \leq \|q - p_j\|^2 \ \forall j \neq i\}. \quad (1)$$

The sensor located at p_i covers all points in V_i . For given sensor locations P , the coverage cost or metric $H(P)$ is given by:

$$H(P) = \sum_{i=1}^k \int_{V_i} \|q - p_i\|^2 \phi(q) dq. \quad (2)$$

The optimal sensor coverage problem is to find an optimal sensor configuration $P^* = (p_1^*, p_2^*, \dots, p_k^*)$ such that the coverage metric is minimized.

$$P^* = \arg \min_P H(P). \quad (3)$$

This problem is also referred to as a *continuous k -median problem* [2], [8].

3.2 Lloyd descent for coverage control

In [2] Cortés et al. show that *centroidal Voronoi configuration*, where each sensor is at the centroid of its own Voronoi partition, is a (locally) optimal solution to the above coverage problem. Further, they propose continuous and discrete-time variations of Lloyd's descent algorithm to control movements of mobile sensors to reach a centroidal Voronoi configuration starting from an initial configuration. In these algorithms, a sensor moves towards the center of its Voronoi partition. As the sensors move, Voronoi partitions as well as their centroids evolve. Cortés et al. prove that the sensors converge to an optimal configuration. During the sensor location update, the algorithm ensures that the coverage metric reduces. The authors further propose an asynchronous distributed control scheme and prove its convergence as well. The quality of the optimal solution depends on the initial sensor configuration. There are no guarantees on the locally optimal coverage cost with respect to the global optimum.

3.3 Weighted- D^2 sampling for initial sensor configuration

We propose a weighted- D^2 sampling approach to select initial sensor locations. It consists of two steps. In the first step, we discretize the domain Q using a uniform square grid and select the *centers* of cells as candidate locations. Next we use a special sampling procedure to select k initial sensor locations. As shown in Figure 1, we superimpose a uniform square grid of size $\varepsilon \times \varepsilon$ on Q and create C_1, C_2, \dots, C_n cells. Note that some cells are entire square grid cells, whereas the others on the boundary of Q form convex polygons. We define the *weight* w_i of each cell C_i as follows:

$$w_i = \int_{C_i} \phi(q) dq. \quad (4)$$

Since we assume that $\int_Q \phi(q) dq = 1$,

$$\sum_{i=1}^n w_i = 1. \quad (5)$$

We compute the *center of mass* x_i of each cell C_i as follows:

$$x_i = \frac{1}{w_i} \int_{C_i} q \phi(q). \quad (6)$$

We consider $X = \{x_1, x_2, \dots, x_n\}$ as a set of candidates for the choice of k initial sensor locations. Next we use an iterative sampling procedure to select k locations from X . Let $D(x)$ denote the shortest distance of point x to the sensor locations already chosen.

- (1) Select the first sensor location p_1 as x_i at random from X with the probability proportional to w_i .
- (2) Select the new location as x_i at random from X with the probability $\frac{w_i D(x_i)^2}{\sum_i w_i D(x_i)^2}$. Update $D(x_i)$.
- (3) Repeat Step (2) until we have chosen k locations.

We refer to the above procedure as weighted- D^2 sampling. It is similar to the D^2 -sampling proposed in [1] except that there is an additional weighing with w_i 's. Once we select k initial sensor locations, we continue to apply the Lloyd descent algorithm. We prove the following theorem which is the central result in this paper.

Theorem 1 *Let $P = (p_1, p_2, \dots, p_k)$ denote the set of k initial sensor locations obtained by applying the above weighted- D^2 sampling procedure. Then,*

$$E[H(P)] \leq 8(\ln k + 2)H(P) + 16\sqrt{2}(\ln k + 2)D\varepsilon + \sum_i J_{C_i|x_i}, \quad (7)$$

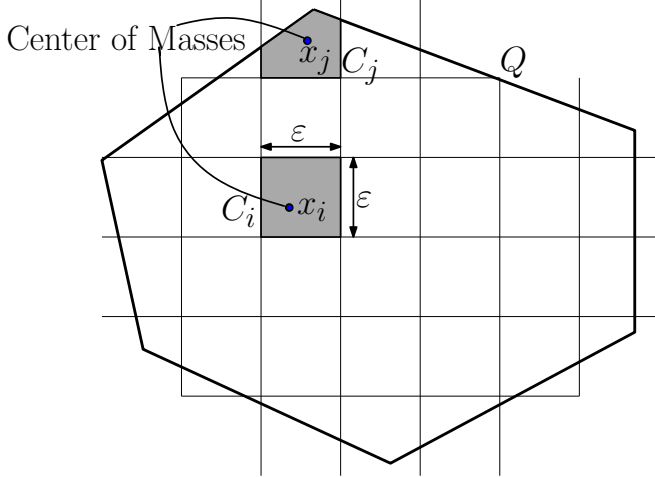


Fig. 1. Superimposition of uniform square grid of size $\varepsilon \times \varepsilon$ on the convex domain Q to generate a cell partition. Cell lying entirely within Q are square grid cells. Cells on the boundary of Q form convex polygonal cells.

where $J_{C_i|x_i}$ is the moment of inertia of cell C_i about its center of mass x_i , and D is the maximum distance between two Voronoi neighbors.

For a sufficiently granular grid division ε the last two terms in the above statement can be ignored and bounded by a small constant. This result implies that the coverage obtained with the weighted- D^2 sampling is $O(\log k)$ -competitive. We prove the above theorem in three parts. In the first part, we establish a relationship between the coverage metric $H(P)$ and the related coverage metric for a weighted k -means problem. In the second part, we show that the weighted- D^2 sampling provides performance guarantees on the coverage metric for the weighted k -means problem. In the third part, we combine the results of the first two parts to prove the main result.

4 Weighted k -means problem and its relation to the coverage problem

We first define the weighted k -means problem in generality. Let $X = \{x_1, x_2, \dots, x_n\}$ be a set of n points in \mathbb{R}^d . Each x_i has an associated non-negative scalar weight $w_i \geq 0$. The weighted k -means problem involves choosing a set of k points $P = (p_1, p_2, \dots, p_k)$ from \mathbb{R}^d such that $\Phi(P) = \sum_{x_i \in X} \min_{p \in P} w_i \|x_i - p\|^2$ is minimized. The difference between the weighted k -means problem and the k -means problem (e.g. [1]) is the weighting by w_i 's.

Now we consider the weighted k -means problem in the context of the coverage problem setting that we discussed in the previous section. Let $X = \{x_1, x_2, \dots, x_n\}$ be the set of the center of masses (Equation 6) of cells C_i 's as we constructed by the grid discretization of Q and let w_i 's be the weights as defined by Equation 4. Now we

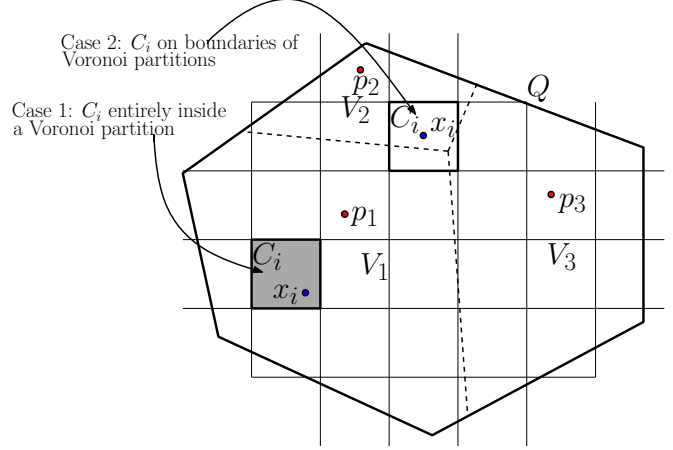


Fig. 2. Two cases based on whether (1) a cell lies entirely inside a Voronoi partition of some sensor location or (2) a cell lies on the boundaries of two or more Voronoi partitions

establish the relationship between the metric $\Phi(P)$ for the weighted k -means problem and the coverage metric $H(P)$ given by the following theorem.

Theorem 2

$$H(P) \leq \Phi(P) + \sum_i J_{C_i|x_i} \leq H(P) + 2\sqrt{2}D\varepsilon, \quad (8)$$

where $J_{C_i|x_i}$ is the moment of inertia of cell C_i about its center of mass x_i , and D is the maximum distance between two Voronoi neighbors.

Proof: Let $h(C_i)$ and $\phi(C_i)$ denote the contributions associated with cell C_i to the metrics $H(P)$ and $\Phi(P)$ respectively. Thus, $H(P) = \sum_i h(C_i)$ and $\Phi(P) = \sum_i \phi(C_i)$, where,

$$h(C_i) = \int_{q \in C_i} \phi(q) \min_{p \in P} (\|q - p\|^2) dq, \quad (9)$$

and,

$$\phi(C_i) = w_i \min_{p \in P} (\|x_i - p\|^2). \quad (10)$$

Now we establish the relationship between $h(C_i)$ and $\phi(C_i)$. As discussed in Equation 1, let V_1, V_2, \dots, V_k denote the Voronoi partition associated with p_1, p_2, \dots, p_k . Now we consider two cases as shown in Figure 3. In the first case, a cell lies entirely within a Voronoi partition for some sensor location. In the second case, a cell lies on the boundaries of two or more Voronoi partitions.

Case 1: Suppose C_i lies entirely within some Voronoi partition. Let V_j denote that partition for point p_j . Then,

$\phi(C_i) = w_i \|x_i - p_j\|^2$ and,

$$h(C_i) = \int_{q \in C_i} \phi(q) \|q - p_j\|^2 dq. \quad (11)$$

By parallel axis theorem,

$$h(C_i) = w_i \|x_i - p_j\|^2 + J_{C_i|x_i}, \quad (12)$$

where $J_{C_i|x_i} = \int_{q \in C_i} \phi(q) \|q - x_i\|^2 dq$. $J_{C_i|x_i}$ is essentially the moment of inertia of cell C_i about its center of mass x_i . Thus,

$$h(C_i) = \phi(C_i) + J_{C_i|x_i}. \quad (13)$$

Case 2: C_i lies on the boundaries of two or more Voronoi partitions. Suppose C_i lies at the boundaries of Voronoi partitions $V_{i_1}, V_{i_2}, \dots, V_{i_m}$ corresponding to the points $p_{i_1}, p_{i_2}, \dots, p_{i_m}$. Without loss of generality, suppose x_i lies in V_{i_1} . Let $D_{i_j} = C_i \cap V_{i_j}$ denote the part of C_i belonging to V_{i_j} . Further, let y_{i_j} denote the center of mass of each D_{i_j} , given by

$$y_{i_j} = \frac{1}{v_{i_j}} \int_{D_{i_j}} q \phi(q) dq, \quad (14)$$

where weight $v_{i_j} = \int_{D_{i_j}} \phi(q) dq$. Note that $w_i = \sum_{j=1}^m v_{i_j}$. Now, $\phi(C_i) = w_i \|x_i - p_{i_1}\|^2$.

$$\begin{aligned} h(C_i) &= \sum_{j=1}^m \int_{q \in D_{i_j}} \phi(q) \|q - p_{i_j}\|^2 dq \\ &= \sum_{j=1}^m \left[v_{i_j} \|y_{i_j} - p_{i_j}\|^2 + J_{D_{i_j}|y_{i_j}} \right] \\ &= \sum_{j=1}^m \left[v_{i_j} \|y_{i_j} - p_{i_j}\|^2 - v_{i_j} \|y_{i_j} - p_{i_1}\|^2 \right] \\ &\quad + \sum_{j=1}^m \left[v_{i_j} \|y_{i_j} - p_{i_1}\|^2 + J_{D_{i_j}|y_{i_j}} \right] \\ &= \sum_{j=1}^m v_{i_j} [\|y_{i_j} - p_{i_j}\|^2 - \|y_{i_j} - p_{i_1}\|^2] \\ &\quad + w_i \|x_i - p_{i_1}\|^2 + J_{C_i|x_i} \\ &= \sum_{j=1}^m v_{i_j} [\|y_{i_j} - p_{i_j}\|^2 - \|y_{i_j} - p_{i_1}\|^2] + \phi(C_i) + J_{C_i|x_i}. \end{aligned}$$

Rearranging the above,

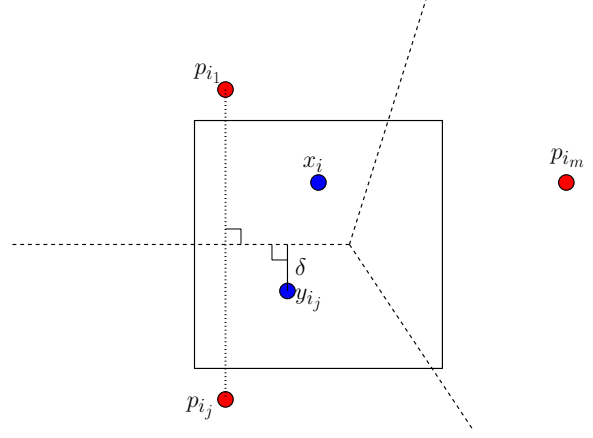


Fig. 3. A detailed illustration of a cell getting divided by two or more Voronoi partitions.

$$h(C_i) + \sum_{j=1}^m v_{i_j} [\|y_{i_j} - p_{i_1}\|^2 - \|y_{i_j} - p_{i_j}\|^2] = \phi(C_i) + J_{C_i|x_i}. \quad (15)$$

Note that each $\|y_{i_j} - p_{i_1}\| \geq \|y_{i_j} - p_{i_j}\|$. Therefore,

$$h(C_i) \leq \phi(C_i) + J_{C_i|x_i}. \quad (16)$$

Now we bound the second term on the left hand side in Equation 15 above. As shown in Figure 3, the Voronoi partition is a perpendicular bisector of segment $P_{i_1}P_{i_j}$. Let δ be the length of the perpendicular projection of y_{i_j} on the Voronoi partition crossing $P_{i_1}P_{i_j}$, which is the perpendicular bisector of $P_{i_1}P_{i_j}$. Suppose $|P_{i_1}P_{i_j}| = \|p_{i_1} - p_{i_j}\| = d_{ij}$.

$$\begin{aligned} \|y_{i_j} - p_{i_1}\|^2 - \|y_{i_j} - p_{i_j}\|^2 &= \left(\frac{d_{ij}}{2} + \delta \right)^2 - \left(\frac{d_{ij}}{2} - \delta \right)^2 \\ &= 2d_{ij}\delta \end{aligned}$$

Note that $\delta \leq \sqrt{2}\varepsilon$. Let D denote the distance between the maximally separated Voronoi neighbors. Thus, $\|p_{i_1} - p_{i_j}\| = d_{ij} \leq D$. Substituting these bounds in the above equation, we obtain,

$$\|y_{i_j} - p_{i_1}\|^2 - \|y_{i_j} - p_{i_j}\|^2 = 2d\delta \leq 2\sqrt{2}\varepsilon D. \quad (17)$$

Substituting Equation 17 into Equation 15,

$$\begin{aligned}\phi(C_i) + J_{C_i|x_i} &\leq h(C_i) + \sum_{j=1}^m v_{ij} 2\sqrt{2}\varepsilon D \\ &\leq h(C_i) + 2\sqrt{2}w_i\varepsilon D\end{aligned}\quad (18)$$

Combining Equation 13 in Case 1 and Equation 16 in Case 2, and summing over all cells, we obtain,

$$H(P) \leq \Phi(P) + \sum_i J_{C_i|x_i}. \quad (19)$$

Combining Equation 13 in Case 1 and Equation 18 in Case 2, and summing over all cells, we obtain,

$$\Phi(P) + \sum_i J_{C_i|x_i} \leq H(P) + 2\sqrt{2}D\varepsilon \sum_{i \in \mathcal{B}(V)} w_i, \quad (20)$$

where $\mathcal{B}(V)$ denotes the set of cells that are on the boundaries of the Voronoi partition. Equation 5 implies that $\sum_{i \in \mathcal{B}(V)} w_i \leq 1$. Substituting above, we obtain:

$$\Phi(P) + \sum_i J_{C_i|x_i} \leq H(P) + 2\sqrt{2}D\varepsilon. \quad (21)$$

Equations 19 and 21 imply the main result of the theorem. \square

5 Guarantees on the weighted k -means solutions with weighted- D^2 sampling

In this section, we show that the weighted- D^2 sampling yields $O(\log k)$ -competitive solution to the weighted k -means problem. Note that this result is not just applicable to the specific weighted k -means instance we considered in the previous section, but holds true in general. We essentially extend the result in [1] for D^2 sampling. Let Φ_{OPT} denote an optimal solution to $\Phi(P)$. We prove the following result.

Theorem 3 *The weighted- D^2 sampling procedure in Section 3.3 applied to the weighted k -means problem yields the following guarantee:*

$$\mathbb{E}[\Phi(P)] \leq 8(\ln k + 2)\Phi_{\text{OPT}}. \quad (22)$$

Since the proof of the above theorem exactly follows and extends the steps in [1], we have included the proof in Appendix 8. Now we are ready to prove Theorem 1.

Proof of Theorem 1: Using the left hand side bound in Theorem 2 and Theorem 3,

$$\begin{aligned}\mathbb{E}[H(P)] &\leq \mathbb{E}[\Phi(P)] + \sum_i J_{C_i|x_i} \\ &\leq 8(\ln k + 2)\Phi_{\text{OPT}} + \sum_i J_{C_i|x_i}.\end{aligned}$$

$\Phi_{\text{OPT}} \leq \Phi(P^*)$, where P^* is the optimal solution to $H(P)$ as in Equation 3. Substituting above and also using the right hand side bound in Theorem 2,

$$\begin{aligned}\mathbb{E}[H(P)] &\leq 8(\ln k + 2)\Phi(P^*) + \sum_i J_{C_i|x_i} \\ &\leq 8(\ln k + 2)H(P) + 16\sqrt{2}(\ln k + 2)D\varepsilon + \sum_i J_{C_i|x_i}.\end{aligned}$$

\square

6 Numerical Results

We performed numerical simulations to compare the sensor coverage performance of the initial sensor configuration chosen with the weighted- D^2 sampling to that with the uniform random initial placement. In particular, for these two types of initial configurations, we compare the initial coverage, the final coverage after applying the Lloyd descent, and the distance traveled by sensors during the descent, which is closely related to the energy spent during the descent. We performed our simulations in Matlab. We set the domain Q as a unit square with the vertices at $[0, 0]$, $[1, 0]$, $[1, 1]$ and $[0, 1]$. We consider the density function $\phi(x, y)$ as a superposition of two Gaussians given by:

$$\begin{aligned}\phi(x, y) &= \frac{1}{A}e^{-10(x-0.75)^2 - 2(y-0.75)^2} \\ &\quad + \frac{1}{A}e^{-20(x-0.25)^2 - 2(y-0.25)^2}\end{aligned}\quad (23)$$

The centers of two Gaussians are located at $(0.25, 0.25)$ and $(0.75, 0.75)$. $A = 0.610882$ is the normalization constant such that $\int_Q \phi(x, y) dx dy = 1$. In order to obtain the initial sensor configuration with the weighted- D^2 sampling, we discretize the domain into grid cells of size $\varepsilon \times \varepsilon$ and sample k locations from the set of the center of masses of the grid cells according to the algorithm in Section 3.3. For the case of the initial configuration with the uniform random sampling, we choose k locations uniformly at random from the unit square. We present the results for the three scenarios. In the first scenario, we set $k = 10$ and choose $\varepsilon = 0.1$, i.e., we choose a grid division of 10×10 cells. In the second scenario, we set $k = 10$, but choose $\varepsilon = 0.05$ with 20×20 grid cells to understand the effect of discretization. In the last scenario, we increase the number of sensors to $k = 20$ with $\varepsilon = 0.05$. In each

scenario, we performed 50 simulation runs for each type of the initial sensor configuration. In each run, we apply the Lloyd descent until the sensors converge to the final configuration. In each iteration of the Lloyd descent, every sensor moves towards the center of its Voronoi cell with the proportional gain of $K = 10$. We choose the convergence criterion as the mean $L1$ -norm of the sensor position change during each iteration being less than $1E-4$. During each iteration, we compute the sensor coverage and also keep track of the distances traveled by the sensors. We discuss our observations below.

As an illustration, from the runs for the first scenario of $k = 10$ and $\varepsilon = 0.1$, Figure 4 (a) and (b) respectively show an example initial sensor configuration for the weighted- D^2 sampling case and the uniform random case. In each case, the contours of the underlying field density function $\phi(\cdot)$ are shown. The Voronoi diagrams of the sensor locations are also shown. Figure 5 shows the coverage for each case during the iterations of the Lloyd descent. In this example, the initial coverage for the weighted- D^2 case is much lower than the uniform random case. We consistently observe this trend across multiple runs as we will discuss shortly. In both cases, across the iterations, the final coverage is close to each other. However, for the uniform random deployment, in the initial round of iterations each sensor travels much longer distance. Figure 4 (c) and (d) show the trajectories of the sensor locations in each case during the Lloyd descent. The red squares represent the final sensor configuration after the convergence. We note that the paths traveled in the uniform random case are considerably longer than the weighted- D^2 sampling case, and again, we consistently observe this trend across multiple runs.

In Table 1 and Table 2, we present the summary statistics across multiple runs for each of the three scenarios considered here. Table 1 shows, for the cases of weighted- D^2 and uniform random deployments, the average initial coverage, the average final coverage, the standard deviations across multiple runs and also the % improvement in the initial coverage with the weighted- D^2 sampling over the uniform random sampling. For all the scenarios, both types of deployments lead to similar final coverage values through the Lloyd descent. For each scenario, the initial coverage with the weighted- D^2 sampling is about 1.5 times the final coverage value, whereas for the uniform random case, the initial coverage is 2.3 times the final coverage value. In each scenario, the weighted- D^2 sampling yields on average over 30% improvement in the initial coverage than the uniform random deployment. Between the first and second scenario, we note that the initial and final coverage values for the weighted- D^2 sampling are not too different indicating that the grid size granularity of $\varepsilon = 0.1$ is sufficient compared to $\varepsilon = 0.05$.

Table 2 shows for each case of the deployment the comparison of the average across multiple runs for the average distance traveled by each sensor during the Lloyd

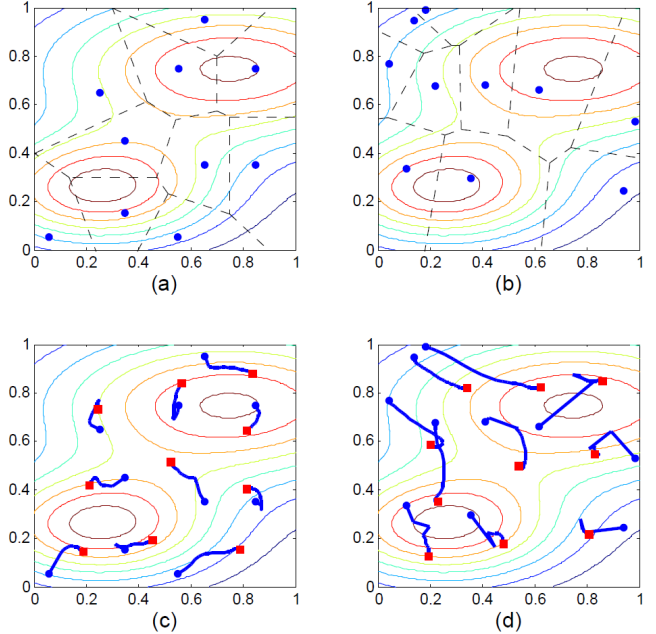


Fig. 4. An illustrative example to show the comparison of the initial sensor configurations and their Voronoi partitions for the weighted- D^2 (Figure a) and the uniform random sampling cases (Figure b). Figures (c) and (d) also show the trajectories of the sensors during the Lloyd descent from the initial sensor configuration (blue circles) to the final configuration (red squares). Figure (c) and (d) show the weighted- D^2 and the uniform random sampling cases respectively.

descent. The table also shows the standard deviations across the runs and the % improvement in the average distance traveled per sensor with the weighted- D^2 sampling over the uniform random case. We note that the average distance traveled per sensor is reduced by at least 30% for $k = 10$ with the weighted- D^2 sampling over the uniform sampling case and the reduction is 25% for $k = 20$. Thus, besides providing the guarantee on the optimal solution before even applying the Lloyd descent, the weighted- D^2 also leads to significant energy savings during the descent. If the sensor motion dynamics is the same in both cases, this also implies faster convergence to the optimal solution.

7 Conclusions and Future work

In this paper we revisited the mobile sensor coverage problem formulated as a continuous locational optimization problem. The Lloyd descent algorithm yields a locally optimal solution with the guaranteed convergence. Researchers have extensively studied variations of this coverage problem and have proposed distributed versions of the variations of the Lloyd descent algorithm to achieve locally optimal solutions with guaranteed convergence. The quality of the final solution depends on the initial sensor configuration which is typically chosen uniformly at random. In this paper, we focus on the original

Case		Weighted- D^2 initial configuration		Uniform random initial configuration		% Improvement
Sensors (k)	Grid size (ε)	Initial coverage (H_{D^2})	Final coverage ($H_{D^2}^*$)	Initial coverage (H_U)	Final coverage (H_U^*)	Initial coverage ($\frac{H_U - H_{D^2}}{H_U}$)%
10	0.1	0.0235 ± 0.0023	0.0154 ± 0.0001	0.0372 ± 0.0085	0.0155 ± 0.0002	36.7
10	0.05	0.0236 ± 0.0024	0.0154 ± 0.0001	0.0353 ± 0.0115	0.0154 ± 0.0001	33.1
20	0.05	0.0121 ± 0.0011	0.0077 ± 0.0000	0.0179 ± 0.0054	0.0077 ± 0.0000	32.4

Table 1

Comparison of the means and the standard deviations of the initial coverage and the final coverage after the Lloyd descent across 50 runs for two types of initial sensor configurations for three scenarios. The last column shows the % improvement in the average initial coverage value for the weighted- D^2 sampling over the uniform random sampling case.

Case		Average distance traveled per sensor during the Lloyd descent		% Improvement
Sensors (k)	Grid size (ε)	Weighted- D^2 initial configuration (d_{D^2})	Uniform random initial configuration (s_U)	Average distance ($\frac{d_U - d_{D^2}}{d_U}$)%
10	0.1	0.2281 ± 0.0512	0.3441 ± 0.0747	33.7
10	0.05	0.2159 ± 0.0485	0.3173 ± 0.0930	32.0
20	0.05	0.1633 ± 0.0287	0.2192 ± 0.0429	25.5

Table 2

Comparison of the means and the standard deviations of the average distance traveled by the sensors during the Lloyd descent across 50 runs for two types of initial sensor configurations for three scenarios. The last column shows the % improvement in the average distance traveled per sensor for the weighted- D^2 sampling over the uniform random sampling case.

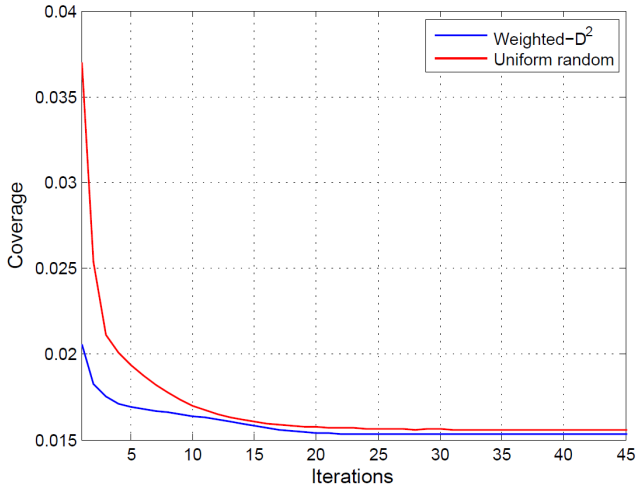


Fig. 5. An illustrative example to show the comparison of the coverage value updates during the Lloyd descent for the weighted- D^2 and the uniform random sampling cases.

coverage problem formulation and propose the weighted- D^2 sampling to choose the initial sensor configuration and show that it yields $O(\log k)$ -competitive sensor coverage before even applying the Lloyd descent. We proved our result in two steps. In the first step, we established a close relationship between the sensor coverage problem and a suitably selected weighted k -means problem. In the second step, we extended the original result in [1] to the weighted- D^2 sampling for the weighted k -means problem. Through extensive numerical simulations, we show that the initial coverage with the weighted- D^2 sam-

pling is significantly lower than the uniform random initial sensor configuration. We also showed that the average distance traveled by the sensors to reach the final configuration through the Lloyd descent is also significantly lower with the weighted- D^2 sampling than that with the uniform random initial deployment. This also means considerable savings in the energy spent by the sensors during motion and faster convergence.

In future, we plan to extend this work to address guaranteed coverage for the variations of the coverage problem such as generic monotonic sensing performance functions. We also plan to extend the sampling procedure to include the learning of the density function.

References

- [1] David Arthur and Sergei Vassilvitskii. k -means++: the advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, SODA '07*, pages 1027–1035, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.
- [2] J. Cortés, S. Martínez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20(2):243–255, 2004.
- [3] Jorge Cortes, Sonia Martinez, and Francesco Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM Control, Optimisation & Calculus of Variations*, 11:691–719, 2005.
- [4] Jorge Cortes, Sonia Martinez, Timur Karatas, and Francesco Bullo. Coverage control for mobile sensing networks. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 2, pages 1327–1332. IEEE, 2002.

- [5] Jason Derenick, Nathan Michael, and Vijay Kumar. Energy-aware coverage control with docking for robot teams. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 3667–3672. IEEE, 2011.
- [6] Ajay Deshpande, Sameera Poduri, Daniela Rus, and G Sukhatme. Distributed coverage control for mobile sensors with location-dependent sensing models. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 2344–2349. IEEE, 2009.
- [7] Ajay Deshpande, Sanjay E Sarma, Kamal Youcef-Toumi, and Samir Mekid. Optimal coverage of an infrastructure network using sensors with distance-decaying sensing quality. *Automatica*, 49(11):3351–3358, 2013.
- [8] Sándor P. Fekete, Joseph S. B. Mitchell, and Karin Weinbrecht. On the continuous weber and k-median problems (extended abstract). In *Proceedings of the Sixteenth Annual Symposium on Computational Geometry, SCG '00*, pages 70–79, New York, NY, USA, 2000. ACM.
- [9] Andrew Kwok and S Martinez. Energy-balancing cooperative strategies for sensor deployment. In *Decision and Control, 2007 46th IEEE Conference on*, pages 6136–6141. IEEE, 2007.
- [10] Atsuyuki Okabe and Atsuo Suzuki. Locational optimization problems solved through voronoi diagrams. *European Journal of Operational Research*, 98(3):445–456, 1997.
- [11] S Salapaka, A Khalak, and MA Dahleh. Constraints on locational optimization problems. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 2, pages 1741–1746. IEEE, 2003.
- [12] Mac Schwager, James McLurkin, and Daniela Rus. Distributed coverage control with sensory feedback for networked robots. In *Robotics: Science and Systems II*, pages 49–56. MIT Press, 2007.
- [13] Mac Schwager, Daniela Rus, and Jean-Jacques Slotine. Decentralized, adaptive coverage control for networked robots. *The International Journal of Robotics Research*, 28(3):357–375, 2009.

8 Appendix: Proof of Theorem 3

Let $P^o = \{p_1^o, p_2^o, \dots, p_k^o\}$ denote an optimal solution to $\Phi(P)$. Thus, $\Phi_{\text{OPT}} = \Phi(P^o)$.

Lemma 4 *Let p_a^o denote an arbitrary center from the optimal solution P^o and let A denote the cluster of x_i 's that are covered by p_a^o . Consider another clustering with just one center x_i which is chosen from A with the probability proportional to w_i . Then,*

$$\mathbb{E}[\phi(A)] = 2\phi_{\text{OPT}}(A). \quad (24)$$

Proof: p_a^o is the center of mass of A . Thus,

$$p_a^o = \frac{\sum_{i \in A} w_i x_i}{\sum_{i \in A} w_i}. \quad (25)$$

$$\begin{aligned} \mathbb{E}[\phi(A)] &= \sum_{i \in A} \frac{w_i}{\sum_{j \in A} w_j} \sum_{j \in A} w_j \|x_i - x_j\|^2 \\ &= \sum_{i \in A} \frac{w_i}{\sum_{j \in A} w_j} \sum_{j \in A} w_j \|x_i - p_a^o - (x_j - p_a^o)\|^2 \\ &= \sum_{i \in A} \frac{w_i}{\sum_{j \in A} w_j} \sum_{j \in A} w_j \|x_i - p_a^o\|^2 + \\ &\quad \sum_{i \in A} \frac{w_i}{\sum_{j \in A} w_j} \sum_{j \in A} w_j \|x_j - p_a^o\|^2 - \\ &\quad \sum_{i \in A} \frac{w_i}{\sum_{j \in A} w_j} \sum_{j \in A} 2w_j (x_i - p_a^o) \cdot (x_j - p_a^o). \end{aligned}$$

In the last equation above, note that the first two terms on the right hand side are the same and each is reduced to $\sum_{i \in A} w_i \|x_i - p_a^o\|^2$. The third term is 0 by the definition of center of mass in Equation 25. Therefore,

$$\mathbb{E}[\phi(A)] = 2 \sum_{i \in A} w_i \|x_i - p_a^o\|^2 = 2\phi_{\text{OPT}}(A).$$

□

Lemma 5 *Let p_a^o denote an arbitrary center from the optimal solution P^o and let A denote the cluster of x_i 's that are covered by p_a^o . Consider an arbitrary clustering \mathcal{C} . If we add another center to \mathcal{C} chosen from A with weighted- D^2 sampling, then $\mathbb{E}[\phi(A)] \leq 8\phi_{\text{OPT}}(A)$.*

Proof: Let $D(x)$ denote the shortest distance of x to the already existing centers in \mathcal{C} . Given that we add a new center from A , the probability that we choose $x_i \in A$ is $\frac{w_i D(x_i)^2}{\sum_{i \in A} w_i D(x_i)^2}$. After choosing x_i , each $x_j \in A$ will contribute $w_j \min(D(x_j), \|x_i - x_j\|)^2$ to $\phi(A)$. Therefore,

$$\mathbb{E}[\phi(A)] = \sum_{i \in A} \frac{w_i D(x_i)^2}{\sum_{i \in A} w_i D(x_i)^2} \sum_{j \in A} w_j \min(D(x_j), \|x_i - x_j\|)^2. \quad (26)$$

By triangle inequality, $D(x_i) \leq D(x_j) + \|x_i - x_j\|$ for all x_i, x_j . By squaring both sides, we have $D(x_i)^2 \leq D(x_j)^2 + \|x_i - x_j\|^2 + 2\|x_i - x_j\|D(x_j)$. Note that $D(x_j)^2 + \|x_i - x_j\|^2 \geq 2\|x_i - x_j\|D(x_j)$. Therefore we have, $D(x_i)^2 \leq 2D(x_j)^2 + 2\|x_i - x_j\|^2$. Multiplying by $\frac{w_j}{\sum_{j \in A} w_j}$ and summing over all $j \in A$, we get,

$$\begin{aligned} D(x_i)^2 &\leq \frac{2}{\sum_{j \in A} w_j} \sum_{j \in A} w_j D(x_j)^2 \\ &\quad + \frac{2}{\sum_{j \in A} w_j} \sum_{j \in A} w_j \|x_i - x_j\|^2 \end{aligned}$$

Substituting in 26,

$$\begin{aligned} E[\phi(A)] &\leq \sum_{i \in A} \frac{2w_i}{\sum_{j \in A} w_j} \frac{\sum_{j \in A} w_j D(x_j)^2}{\sum_{i \in A} w_i D(x_i)^2} \\ &\quad \times \sum_{j \in A} w_j \min(D(x_j), \|x_i - x_j\|)^2 \\ &\quad + \sum_{i \in A} \frac{2w_i}{\sum_{j \in A} w_j} \frac{\sum_{j \in A} w_j \|x_i - x_j\|^2}{\sum_{i \in A} w_i D(x_i)^2} \\ &\quad \times \sum_{j \in A} w_j \min(D(x_j), \|x_i - x_j\|)^2. \end{aligned}$$

In the first expression, we substitute $\min(D(x_j), \|x_i - x_j\|)^2 \leq \|x_i - x_j\|^2$, and in the second expression, we substitute, $\min(D(x_j), \|x_i - x_j\|)^2 \leq D(x_j)^2$.

$$\begin{aligned} E[\phi(A)] &\leq 4 \sum_{i \in A} \frac{w_i}{\sum_{i \in A} w_i} \sum_{j \in A} w_j \|x_i - x_j\|^2 \\ &= 8\phi_{\text{OPT}}(A). \end{aligned}$$

The last step above follows from Lemma 4. \square

We have now shown that as long as our sampling chooses centers from each cluster from the optimal solution P^o , the cost metric is competitive. The lemma below bounds the total error.

Lemma 6 *Let \mathcal{C} be an arbitrary clustering obtained by choosing some x_i 's as centers. Let \mathcal{C}_{OPT} be the clustering corresponding to the optimal solution P^o . A cluster from \mathcal{C}_{OPT} is "uncovered" if none of the x_i 's in that cluster are part of the centers of \mathcal{C} . Let \mathcal{X}_u denote the set of "uncovered" points in these clusters. Also, let $\mathcal{X}_c = \mathcal{X} - \mathcal{X}_u$. Now suppose we add $t \leq u$ random centers to \mathcal{C} chosen with the weighted- D^2 sampling. Let \mathcal{C}' denote the resulting clustering and let Φ' denote the corresponding cost metric. Then,*

$$E[\Phi'] \leq (\Phi(\mathcal{X}_c) + 8\Phi_{\text{OPT}}(\mathcal{X}_u)) \cdot (1 + H_t) + \frac{u-t}{u} \cdot \Phi(\mathcal{X}_u).$$

Here, H_t denotes the harmonic sum, $1 + \frac{1}{2} + \dots + \frac{1}{t}$.

The proof of the above lemma follows the exact argument in [1] word to word and hence, it is not repeated here. Now we are ready to prove Theorem 3.

Proof of Theorem 3: Let \mathcal{C} denote the clustering after choosing the first center in Step (1) of the weighted- D^2 sampling. Let A be the cluster from \mathcal{C}_{OPT} which contains this first center. After applying Lemma 6 with $t = u = k - 1$ and A being the only covered cluster, we obtain,

$$E[\Phi] \leq (\Phi(A) + 8\Phi_{\text{OPT}} - 8\Phi_{\text{OPT}}(A)) \cdot (1 + H_{k-1}).$$

The result follows from Lemma 4 and $H_{k-1} \leq \ln k + 1$.